

# PATENT ABSTRACTS OF JAPAN

(11)Publication number : 2000-148864

(43)Date of publication of application : 30.05.2000

(51)Int.Cl. G06F 17/60  
G06F 13/00

(21)Application number : 11-286140 (71)Applicant : INTERNATL BUSINESS MACH CORP <IBM>

(22)Date of filing : 07.10.1999 (72)Inventor : CHARU C AGAAWARU  
PHILIP S YU

(30)Priority

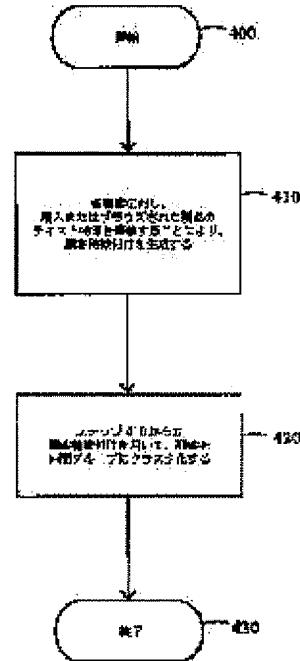
Priority number : 98 169029 Priority date : 09.10.1998 Priority country : US

## (54) METHOD AND DEVICE FOR PROVIDING PRODUCT RECOMMENDATION

(57)Abstract:

**PROBLEM TO BE SOLVED:** To provide a system and method for generating product recommendation at an electronic commercial transaction site by characterizing products and user behavior.

**SOLUTION:** The characterization of a supplied product is led out at a commercial transaction site based on a supplied text description at the site. Customer characterization is made to a customer browsing at the site. The characterizations are related to the product purchased or browsed by the customer and include a set of led-out product characterizations. A group of fellow customers is formed by clustering customers having similar customer characterizations. Then a product is recommended to the customers based on the processed characterizations and the data about the group of fellow customers.





(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開2000-148864

(P2000-148864A)

(43) 公開日 平成12年5月30日 (2000.5.30)

(51) Int.Cl.<sup>7</sup>

G 06 F 17/60  
13/00

識別記号

355

F I

G 06 F 15/21  
13/00

テマコード(参考)

Z  
355

審査請求 有 請求項の数22 OL (全13頁)

(21) 出願番号 特願平11-286140

(22) 出願日 平成11年10月7日 (1999.10.7)

(31) 優先権主張番号 09/169029

(32) 優先日 平成10年10月9日 (1998.10.9)

(33) 優先権主張国 米国 (U.S.)

(71) 出願人 390009531

インターナショナル・ビジネス・マシーンズ・コーポレーション  
INTERNATIONAL BUSINESSES MACHINES CORPORATION  
アメリカ合衆国10504、ニューヨーク州  
アーモンク (番地なし)

(74) 代理人 100086243

弁理士 坂口 博 (外1名)

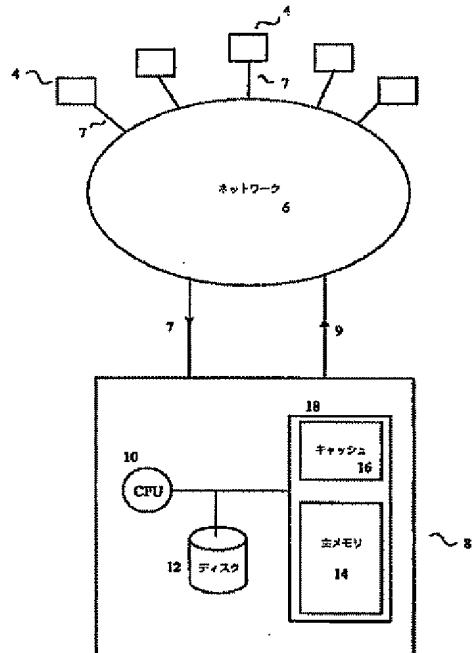
最終頁に続く

(54) 【発明の名称】 製品推薦を提供する方法および装置

(57) 【要約】

【課題】 製品及びユーザ振舞いの特徴付けにより、電子商取引サイトにおいて製品推薦を生成するシステム及び方法を提供すること。

【解決手段】 本発明は、電子商取引サイトにおいて提供される製品のテキスト記述にもとづき、サイトにおいて提供される製品の製品特徴付けを導出する。電子商取引サイトをブラウズする顧客に対して、顧客特徴付けが生成される。特徴付けは、その顧客により購入またはブラウズされた製品に関連付けられ、導出された製品特徴付けの集合を含む。類似の顧客特徴付けを有する顧客をクラスタ化することにより、仲間グループが形成される。次に、処理された特徴付け及び仲間グループ・データにもとづき、顧客に製品を推薦する。



## 【特許請求の範囲】

【請求項1】電子商取引環境において、顧客に製品推薦を提供する方法であって、  
 a) 複数の製品の各々に対する製品特徴付けを導出するステップと、  
 b) 各それぞれの前記顧客による前記製品特徴付けの利用にもとづき、前記顧客の各々に関する、個々の顧客特徴付けを生成するステップと、  
 c) 前記顧客特徴付けの類似点にもとづきクラスタ化し、仲間グループを形成するステップと、  
 d) 個々の前記顧客を前記仲間グループの1つに分類するステップと、  
 e) 前記顧客特徴付け及び前記分類済み仲間グループからの情報にもとづき、前記顧客に製品推薦を生成するステップとを含む、方法。

【請求項2】前記顧客特徴付けを生成するステップが、前記顧客が前記製品をブラウズまたは購入する際、前記製品特徴付けを抽出するステップを含む、請求項1記載の方法。

【請求項3】前記顧客特徴付けを生成するステップが、個々の前記顧客によりブラウズまたは購入された全ての前記製品の各前記製品特徴付けを連結するステップを含む、請求項2記載の方法。

【請求項4】前記製品特徴付けを導出するステップが、前記複数の製品に関するテキスト特徴付けを使用する、請求項1記載の方法。

【請求項5】前記製品特徴付けを生成するステップが、  
 a) 前記テキスト記述内の各単語の発生頻度を見い出すステップと、  
 b) 各前記単語の合計頻度を、全ての顧客における当該前記単語の発生頻度により除算するステップと、  
 c) 各前記単語の標準偏差を見い出すステップと、  
 d) より大きな標準偏差を有する単語を選択するステップと、  
 e) 選択された前記単語にもとづき、製品特徴付けを表現するステップとを含む、請求項1記載の方法。

【請求項6】a) 現オンライン・セッションにもとづき、前記製品特徴付けを導出するステップと、  
 b) メモリから履歴製品特徴付けをアクセスするステップと、

c) ステップa) 及びb) からの前記特徴付けを重み付け連結することにより、前記顧客特徴付けを生成するステップと、  
 d) 前記仲間グループの各々のクラスタ中心を計算するステップと、  
 e) ステップc) で生成された前記顧客特徴付けに最も近い前記クラスタ中心を有する前記仲間グループを選択するステップと、  
 f) 選択された前記仲間グループにもとづき、製品推薦、仲間グループ推薦及びプロファイル推薦の1つを生

成するステップとを含む、請求項1記載の方法。

【請求項7】前記クラスタ中心が、各前記仲間グループ内の全ての前記顧客特徴付けのテキスト特徴付けを連結することにより計算される、請求項6記載の方法。

【請求項8】前記推薦が、前記現オンライン・セッションにおいて購入及びブラウズされた製品の、テキスト特徴付けの重み付け連結を含む、請求項6記載の方法。

【請求項9】電子商取引環境において、顧客に製品推薦を提供するプロセス・ステップを実行する、コンピュータにより実行可能な命令プログラムを実現する、コンピュータにより読み取り可能な記憶プログラム装置であって、前記命令プログラムが、電子商取引において売りに出される製品に関する製品情報をアクセスするステップと、

前記製品情報から製品特徴付けを導出するステップと、前記製品に関する顧客利用情報をアクセスするステップと、

アクセスされた前記利用情報にもとづき、前記顧客の各々に関する顧客特徴付けを生成するステップと、

前記顧客特徴付けの類似点にもとづきクラスタ化し、複数の仲間グループを形成するステップと、

現セッション顧客から照会を受信するステップと、前記現セッション顧客に関する連付けられる現顧客特徴付けを生成するステップと、

前記現セッション顧客を、前記現顧客特徴付け及び過去の顧客特徴付けの類似点にもとづき、選択された前記仲間グループに分類するステップと、

前記照会に対して、少なくとも分類された選択済みの前記仲間グループからの情報にもとづき、前記現セッション顧客に製品推薦により応答するステップとを含む、装置。

【請求項10】前記コンピュータに、前記現セッション顧客と電子商取引サイト間で交換されるデータを抽出し、前記現顧客特徴付けを生成するように指示する記憶プログラムを含む、請求項9記載の装置。

【請求項11】前記現顧客特徴付けを生成するステップが、前記現セッション顧客によりブラウズまたは購入された全ての製品の前記製品特徴付けの各々を重み付け連結するステップを含む、請求項9記載の装置。

【請求項12】各前記仲間グループ内の全ての前記顧客特徴付けのテキスト特徴付けを連結することにより、当該仲間グループのクラスタ中心を計算するステップと、前記現顧客特徴付けに最も近い前記クラスタ中心を有する前記仲間グループを選択するステップと、

選択された前記仲間グループからのデータにもとづき、製品推薦、仲間グループ推薦及びプロファイル推薦の1つを生成するステップとを含む、請求項9記載の装置。

【請求項13】電子商取引環境において顧客と対話し、顧客照会に応答を提供するようにコンピュータに指示する、プロセッサ及び記憶プログラムを有するコンピュー

タであって、該コンピュータが、複数の製品の製品特徴付けを導出する手段と、前記顧客による前記製品特徴付けの利用にもとづき、前記顧客特徴付けを生成する手段と、前記顧客特徴付けの類似点にもとづきクラスタ化し、複数の仲間グループを形成する手段と、個々の前記顧客を前記仲間グループの1つに分類する手段と、前記顧客特徴付け及び前記仲間グループからの情報にもとづき、前記顧客照会に応答を提供する手段とを含む、記憶プログラム。

【請求項14】前記顧客特徴付けを生成する手段が、前記顧客特徴付けを保管する記憶装置を含む、請求項13記載の記憶プログラム。

【請求項15】現オンライン顧客による前記製品特徴付けの利用から抽出されるデータの連結にもとづき、前記現オンライン顧客の現顧客特徴付けを生成する手段を含む、請求項13記載の記憶プログラム。

【請求項16】前記現顧客特徴付け及び記憶済みの特徴付けの類似点にもとづき、前記現オンライン顧客を前記仲間グループの1つに分類する手段を含む、請求項15記載の記憶プログラム。

【請求項17】前記製品特徴付けを導出する手段が、前記製品に関する付けられるテキスト記述を抽出する手段を含む、請求項13記載の記憶プログラム。

【請求項18】前記応答を提供する手段が、現セッションにおいて個々の前記顧客によりブラウズされた製品からの前記製品特徴付けにもとづき、個々の前記顧客に前記製品を推薦する手段を含む、請求項13記載の記憶プログラム。

【請求項19】前記応答を提供する手段が、履歴製品特徴付けの重み付け頻度、及び現セッションにおいて個々の前記顧客によりブラウズされた製品からの、前記現製品特徴付けの重み付け頻度にもとづき、個々の前記顧客に前記製品を推薦する手段を含む、請求項13記載の記憶プログラム。

【請求項20】前記応答を提供する手段が、履歴製品特徴付けにもとづき、個々の前記顧客に製品を推薦する手段を含む、請求項13記載の記憶プログラム。

【請求項21】前記応答を提供する手段が、プロモーション・リスト内の製品にもとづき、個々の前記顧客に製品を推薦する手段を含む、請求項13記載の記憶プログラム。

【請求項22】前記応答を提供する手段が、前記仲間グループの1つ内の製品頻度にもとづき、仲間分布プロファイルを提供する手段を含む、請求項13記載の記憶プログラム。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】本発明は、電子商取引ネット

ワークを介し、製品推薦を生成するための自動化コンピュータ・ベースの装置及び方法に関し、特に、顧客ユーザの振舞いにもとづき、製品推薦を生成する装置及び方法に関する。

#### 【0002】

【従来の技術】最近のインターネットを介するオンライン・ショッピングの人気の上昇により、ショッピング・サイトを提供するエンティティは、彼らの製品の販売支援のために、買い物客に関する情報の獲得に关心を寄せている。こうした情報を獲得するための従来の顧客または市場調査は、供給者のために適用可能で有用である。Likeminds社 ([www.likeminds.com](http://www.likeminds.com)) 及びFirefly Network社 ([www.firefly.com](http://www.firefly.com))などの会社は、顧客による明示的な格付けにもとづく調査情報を提供し、これは一般に、“協調フィルタリング (collaborative filtering)”を使用する推薦エンジンと呼ばれる。格付けにもとづくエンジンの使用は、一様で特定のタイプの製品に特定の適用性を有する。例えば、Likeminds社の場合、顧客は彼らの好き嫌いの度合いにもとづき、コンパクト・ディスクなどの製品の嗜好の格付けを提供するよう求められる。これらの格付けが次に収集され、後の使用のために保管される。将来のある時点で、他の顧客の以前に保管されたデータにもとづき、新たな顧客に対して製品推薦が生成される。顧客による過去の明示的な格付けにもとづき生成される推薦は、“協調フィルタリング”として知られる。

【0003】協調フィルタリング・アプローチは、顧客が製品の明示的な格付けに参加しない場合、うまく機能しない。電子商取引環境における顧客または買手は、一般に、自分のオンライン時間を最小化したいと思い、オンラインか否かを問わず、特に製品の格付けのために、余分な時間を消費たくない。

【0004】他の会社は、電子商取引ウェブサイトからの抽出テキスト及び情報を用いる“内容ベースの”フィルタリングを提案している。

【0005】例えば、Autonomy社 ([www.autonomy.com](http://www.autonomy.com))により提案される知能インフラストラクチャが挙げられる。このシステムは、インターネット情報提供者のためのスケーラブル内容個人化及び編成エンジンである、エージェントウェア (Agentware) 内容サーバを提供する。この技術は、文書及びウェブサイトから主要概念を抽出し、情報の分類、相互参照、ハイパリンク、及び表現を自動化する。このソフトウェアの顧客プロファイリング・システムは、情報及びサービスの提供者が顧客の関心を理解し、個人化情報を配達することを可能にする。

【0006】知能サーバを提供する別の会社は、Aptex Software社 ([www.aptex.com](http://www.aptex.com))である。Aptex Software社は、テキスト及び他の非構造化内容を自動的に分析し、知能的判断及び推薦を行う“内容マイニング (Content

nt Mining) "法を使用する。

【0007】更に別の会社として、Net Perceptions社 ([www.netperceptions.com](http://www.netperceptions.com)) は、製品の暗黙的または明示的格付けを使用し (Grouplen) 、推薦を提供する。暗黙的格付けは、顧客により購入またはブラウズされた製品のセットを指示す。

【0008】上述の知能サーバの提供及び可用性にも関わらず、明示的なグループ製品格付けに頼ることなく、顧客の自然のブラウジングまたは購入振舞いにもとづき、顧客及び製品の特徴付けを容易にし、仲間のグループ分類にもとづき推薦を提供し、製品推薦プロセスに、本質的な顧客個人化をもたらす方法またはシステムが必要とされる。

【0009】

【発明が解決しようとする課題】従って、本発明の目的は、ユーザ・ブラウジングまたは購入振舞いを含む、製品及びユーザ振舞いの特徴付けにより、電子商取引サイトにおいて、製品推薦を生成するシステム及び方法を提供することである。

【0010】本発明の別の目的は、特徴付けを用いて仲間グループを生成し、ユーザへの推薦を個人化することである。ここで仲間グループとは、以前に製品嗜好が保管されており、新たな顧客のものに類似の製品嗜好パターンを示す顧客の集合である。

【0011】

【課題を解決するための手段】上述の目的が、電子商取引環境において、顧客に製品推薦を提供する本発明による方法により達成される。本方法は、複数の製品の各々に対する製品特徴付けを導出するステップと、各それぞれの顧客による製品特徴付けの利用にもとづき、顧客の各々に関する、個々の顧客特徴付けを生成するステップと、顧客特徴付けの類似点にもとづきクラスタ化し、仲間グループを形成するステップと、個々の顧客を仲間グループの1つに分類するステップと、顧客特徴付け及び分類済み仲間グループにもとづき、顧客に製品推薦を生成するステップとを含む。

【0012】顧客特徴付けを生成するステップは、好適には、顧客が製品をブラウズまたは購入する際、製品特徴付けを抽出するステップを含む。また、顧客特徴付けを生成するステップは、個々の顧客によりブラウズまたは購入された全ての製品の各製品特徴付けを連結するステップを含み得る。

【0013】好適には、製品特徴付けは各製品のテキスト特徴付けから導出され、更に、テキスト記述内での各単語の発生頻度を見い出すステップと、各単語の合計頻度を、全ての顧客におけるその単語の発生頻度により除算するステップと、各単語の標準偏差を見い出すステップと、より大きな標準偏差を有する単語を選択するステップと、選択された単語にもとづき、製品特徴付けを表現するステップとを含む。

【0014】本発明の別の方針は、好適には、a) 現オンライン・セッションにもとづき、製品特徴付けを導出するステップと、b) メモリから履歴製品特徴付けをアクセスするステップと、c) ステップa) 及びb) からの特徴付けを重み付け連結することにより、顧客特徴付けを生成するステップと、d) 仲間グループの各々のクラスタ中心を計算するステップと、e) ステップc) で生成された顧客特徴付けに最も近いクラスタ中心を有する仲間グループを選択するステップと、f) 選択された仲間グループにもとづき、製品推薦、仲間グループ推薦及びプロファイル推薦の1つを生成するステップとを含む。クラスタ中心は、各仲間グループ内の全ての顧客特徴付けのテキスト特徴付けを連結することにより計算され、推薦は、現オンライン・セッションにおいて購入及びブラウズされた製品の、テキスト特徴付けの重み付け連結を含み得る。

【0015】本発明によるシステムは、電子商取引環境において顧客と対話し、顧客照会に応答を提供するようにコンピュータに指示するプロセッサ及び記憶プログラムを有するコンピュータを提供する。記憶プログラムは、複数の製品の製品特徴付けを導出する手段と、顧客による製品特徴付けの利用にもとづき、顧客特徴付けを生成する手段と、顧客特徴付けの類似点にもとづきクラスタ化し、複数の仲間グループを形成する手段と、個々の顧客を仲間グループの1つに分類する手段と、顧客特徴付け及び仲間グループからの情報にもとづき、顧客照会に応答を提供する手段とを含む。

【0016】システムは好適には、顧客特徴付けを保管する記憶装置を含む。

【0017】記憶プログラムは更に、現オンライン顧客のために、その顧客による製品特徴付けの利用から抽出されるデータの連結にもとづき、現顧客特徴付けを生成する手段を含む。

【0018】記憶プログラムは更に、現顧客特徴付け及び記憶済みの特徴付けの類似点にもとづき、現オンライン顧客を仲間グループの1つに分類する手段を含む。

【0019】製品特徴付けを導出する手段は、好適には、製品に関連付けられるテキスト記述を抽出する手段を含む。

【0020】応答を提供する手段は、好適には、現セッションにおいて個々の顧客によりブラウズされた製品からの製品特徴付けにもとづき、個々の顧客に製品を推薦する手段を含む。

【0021】応答を提供する手段は、好適には、履歴製品特徴付けの重み付け頻度、及び現セッションにおいて個々の顧客によりブラウズされた製品からの、現製品特徴付けの重み付け頻度にもとづき、個々の顧客に製品を推薦する手段を含む。

【0022】応答を提供する手段はまた、好適には、履歴製品特徴付けにもとづき、個々の顧客に製品を推薦す

る手段を含む。

【0023】

【発明の実施の形態】本発明は一般に、事前処理ステージと、それに続くオンライン照会ステージの2つのステージを含むように特徴化される。事前処理ステージは一般に、電子商取引サイトにおいて入手可能な製品のテキスト記述を簡潔な製品特徴付けに変換するステップを含む。本開示の目的上、用語"特徴付け (characterization)"は、適切な重みを有する単語のセットとして述べられる。重みは、製品のテキスト記述内の単語の発生頻度により定義される。特徴付けは、顧客の振舞いの指標として使用され得る。例えば、顧客に映画が推薦されている応用例では、貴重な製品特徴付けは、映画のクラスと共に、男優及び女優を記述するテキストを示し得る。

【0024】一旦製品特徴付けが生成されると、次に個々の顧客特徴付けが生成される。顧客特徴付けは、顧客により購入またはブラウズされた全ての製品の製品特徴付けの各々を連結することにより生成される。本発明の目的上、連結は単語及び重みのセットを追加し、より長い特徴付けを形成するプロセスとして述べられる。例えば、特徴付け1、すなわちヘビー(1)、メタル(1)、SONY(2)と、特徴付け2、すなわちSONY(1)、音楽(3)が与えられる場合、2つの特徴付けの連結は、次の特徴付け、すなわち、連結特徴付けヘビー(1)、メタル(1)、SONY(3)、音楽(3)を提供する。ここで括弧内の数字は重みに対応する。

【0025】次に、顧客の仲間グループが、顧客特徴付けをクラスタ化することにより形成される。一般に、同趣向の購入またはブラウジング習慣を示す顧客は、仲間グループにクラスタ化される。一旦クラスタが生成されると、事前処理ステージが完了する。事前処理ステージは、所定の時間間隔で実行され、テキスト特徴付け及びクラスタを生成する。クラスタは、オンライン照会ステージにおいて提起される照会に応答するために使用される。

【0026】オンライン照会ステージは事前処理ステージの後に続き、一般に、1人以上の顧客がネットワークを介して、サーバに照会を生成するステップを含む。サーバは照会の結果を計算し、それらを各ぞれの顧客に返却する。

【0027】図面を参照すると、図1は、本発明の実施例に従う好適なアーキテクチャを示す。サーバ8はネットワーク6を介して、複数の顧客またはクライアント4に接続される。クライアントはサーバ8に要求(照会)7を生成し、サーバ8は各照会に対する結果9を返却することにより、要求に応答する。ネットワーク6は、製品販売及び購入を容易にする電子商取引ネットワークである。このネットワークは好適にはインターネットである。

【0028】サーバ8は、サーバを調整及び制御するCPU10を含み、照会結果9を生成するために必要な計算を実行する。主メモリ14は、各顧客4のためのデータ記憶部として作用する。顧客データは、ディスク12上またはキャッシュ16内にも存在し得る。顧客特徴付けが履歴データから導出される場合、その履歴データがその顧客のために、データ記憶部から検索される。

【0029】図2を参照すると、本発明の全体動作の例を示すフロー図が示され、ステップ210及び220を含む事前処理ステージと、ステップ230、240及び250を含むオンライン照会ステージとを有する。

【0030】ステップ210で、電子商取引において各製品に関連付けられるテキスト記述が、製品特徴付けに変換される。本開示の目的上、製品特徴付けは、無関係な用語(すなわち共通言語)を除去するためにフィルタリングされたテキスト特徴付けである。テキスト記述を製品特徴付けに変換するプロセスは、一般に、電子商取引サイトにおいて提供される、各製品のテキスト記述内に含まれる単語を選択するステップを含み、これらの単語は、顧客の振舞いの貴重な指標であることが判明している。例えば、映画またはCDSが推薦されるサイトにおいて、高い推理性を伴う幾つかの単語の例は、"アクション"、"ロマンス"、"ミステリ"、及び"ドラマ"などである。なぜなら、これらの単語は顧客の振舞いを高度に示し、異なる顧客の間で、その分布にかなりの偏りを有しがちであるからである。製品特徴付けの生成に関連するステップは、図3に関連して以下で更に詳述される。

【0031】次にステップ220で、製品特徴付けを用いて、顧客をクラスタまたは仲間グループにグループ化する。仲間グループは、推薦が生成される顧客の購入またはブラウジング特性によく合致する特性を有する、顧客のセットまたはクラスタを構成する。クラスタ化アルゴリズムは公知である。好適には、1998年4月30日付けのC. Aggarwallaによる米国特許出願第070600号"System and Method for Detecting Clusters of Information"で述べられるクラスタ化方法が、クラスタ化操作を実行するために使用される。

【0032】オンライン照会ステージがステップ230で開始し、1人以上の顧客からの照会が受信されて処理される。顧客要求を受信するプロセスは、顧客がいつでもまたはランダム間隔で要求をサーバに発行し得るという点で、本質的に連続的と見なされる。顧客要求がステップ230で受信される。事前処理ステージで生成された事前処理データ(すなわち仲間グループまたはクラスタの形成)にもとづき、顧客照会に対する応答が、ステップ240で生成される。

【0033】ステップ250で、事前処理ステージが最後に実行されて以来、所定の時間間隔が経過したか否かが判断される。所定の時間間隔が経過していない場合、

プロセスはステップ230乃至250のループ内に留まり、追加のユーザ照会がキュー内に存在することに応答する。所定の時間間隔が経過した場合、プロセスはステップ210に戻り、事前処理ステージを実行する。このタイミング構造は、顧客照会を“オンザフライ式”に処理することを容易にする。事前処理ステージは照会ステージに比較して、より多くの時間を消費するので、システムは好適には、事前処理をより低頻度に、しかしながらオンライン照会が低トラフィックに遭遇しているときは常時、実行するように設定される。

【0034】次に図3を参照すると、本発明の1実施例に従い、テキスト記述から製品特徴付けを生成するステップ210(図2参照)のフロー図が示される。ステップ310で、プロセスは、各単語が各顧客により購入された平均頻度を決定する。本発明の目的上、単語は、それが顧客により購入またはブラウズされたある製品のテキスト特徴付け内で発生する場合、顧客により“購入さ\*

$$\text{標準偏差 (単語)} = \sqrt{\frac{1}{N} \sum_{\text{1}}^N (\text{F}(\text{単語, 顧客}) - 1/N)^2} / N^{-1/2}$$

【0035】一般に、低い推理力を有するテキスト記述(すなわち非記述的)からの単語は、言語の総称的な単語(すなわちis, the, an)を含み、低い標準偏差値を有すると思われる。それに対して高い推理力を有する単語(例えば、a singer=\$ name、またはCD音楽製品の推薦環境の一部としての音楽カテゴリなどの単語)は、高い標準偏差値を有すると思われる。最も高い標準偏差を有する単語が、製品の入力テキスト記述から製品特徴付けを生成するために選択される。

【0036】例えば、コンパクト・ディスクの場合の原テキスト記述は、次のようにある。

CD名: Whitford/St. Holmes - Whitford/St. Holmes  
ラベル: SONY MUSIC

ジャンル: 一般

原発売日: 1981

エンジニア: George Pappas

プロデューサ: Tom Allom

ディスク番号: 1

モノ/ステレオ: ステレオ

スタジオ/ライブ: スタジオ

カテゴリ: ヘビー・メタル

【0037】図3のフロー図に示されるプロシージャを実行後、特徴化された単語は、ヘビー、メタル、一般、SONY、Tom、Allomとなる。なぜなら、これらの単語は、CDの性質を高度に示すからである。一方、残りの単語は単に一般的な単語である。日付やラベルなどの単語は、各々の及びあらゆるテキスト記述内に発生しがちであり、従ってこれらの分布は偏りを有さず、全ての人々が“趣味(taste)”に関係無しに、これらの単語をブラウズし得る。他方、“ヘビー”、“メタル”

\*れた”と見なされる。例えば、同一の製品がk回購入された場合、挨拶語(complimentary word)もk回カウントされ、その単語の総頻度はkである。他の実施例では、“購入された”単語を顧客により、またはブラウジング及び購入の予め定義された組み合わせにより、ブラウズされたテキストまたは単語と見なす。顧客により購入された各単語はカウントされ、総頻度として合計される。ステップ320で、各顧客に対するカウントを、各単語が全ての顧客により“購入”された総頻度により除算することにより、総頻度カウントが分数に変換され、F(単語、顧客)として定義される。ステップ330で、異なる顧客により購入された単語の分数の標準偏差が計算される。所与の顧客により購入された単語の分数の標準偏差は、Nを顧客の総数とすると、次のように計算される。

【数1】

などの単語の場合、これらの単語を頻繁にアクセスする人々に著しく偏りがちになる。

【0038】次に図4を参照すると、生成された製品特徴付けから、クラスタ(仲間グループ)を生成する、図2のステップ220に関連付けられるステップのフロー図が示される。ステップ410で、顧客が購入またはブラウズした全ての製品に対する製品特徴付けを連結することにより、顧客特徴付けが生成される。現文脈における連結とは、顧客により購入またはブラウズされた個々の製品特徴付けの各々が附加され、1つの長い特徴付けを形成することを意味する。ステップ420で、各顧客が前のステップで形成された顧客特徴付けにもとづき、仲間グループにクラスタ化される。ここで仲間グループとは、類似の購入またはブラウジング振舞いを示す顧客のセットを定義する。例えば、次の単語、すなわちチップ、ハードウェア、PC、ソフトウェアゲーム、PCゲーム、ジョイスティック、プロセッサ、アブティバ、ウインドウズ98などにより生成される製品特徴付けについて考えてみよう。繰り返し発生する3つの単語“ソフトウェアゲーム、PCゲーム、及びジョイスティック”を有する顧客のクラスタでは、このクラスタの顧客は、コンピュータ・ゲームに关心がある人々のグループに対応することがわかる。ジョイスティック、PCゲーム、及びソフトウェアゲームは、個々の製品の名前ではなく、製品を特徴付ける単語であっても、このクラスタに属するこの単語セットに含まれる情報量は振舞い情報であり、このクラスタの人々に関する特定の仮定が成され得る。

【0039】次に図5を参照すると、顧客要求に対応する様々な照会に、プロセスが応答する(図2のステップ

240) 様子を示すフロー図が示される。図5は、オンライン照会ステージを定義する。図5に示されるプロセス・ステップの各々は、1つの顧客要求に関して述べられる。しかしながら、本方法が到来順に、1人以上の顧客からの複数の要求にも適応し得ることが理解できよう。

【0040】ステップ502で、前に生成された全てのクラスタまたは仲間グループを含む第1の入力が、システムにより受信される。次にステップ505で、第2及び第3の入力、例えば第2の入力は購入振舞い及び第3の入力はブラウジング振舞いが、プロセスにより受信される。第2の入力は、現オンライン購入またはブラウジング・セッションにおいて生成された、製品特徴付けを含むように定義される。製品特徴付けが、図3で示したように生成され得る。好適には、この生成は、顧客が現オンライン・セッションにおいて製品をブラウズまたは購入する間に、リアルタイムに実行される。第3の入力の製品特徴付けは、第2の入力に類似であるが、この製品特徴付けは、主メモリ14に記憶された履歴保管データから取り出される。生成されたデータは、その顧客に特定の以前のブラウジングまたは購入オンライン・セッションからの製品特徴付けを含む。ステップ510で、その顧客に対する顧客特徴付けが生成される。好適には、顧客特徴付けは、第1の入力及びクラスタと、第2及び第3の入力のいずれかまたは両方を用いて構成される。第2及び第3の入力の両方が使用される場合、それらは重み付け連結により結合される。重み付け連結は、第1及び第2の入力からの各単語の頻度を、適切な重みにより乗じることにより導出される。CDの連結における上述の例の場合、特徴付けAがヘビー(1)、メタル(1)、SONY(2)であり、特徴付けBがSONY(1)、音楽(3)である。連結Aに2が乗じられ、連結Bに3が乗じられると、重み付け連結はその製品の総和であり、ヘビー(2)、メタル(2)、SONY(7)、及び音楽(9)を生成する。重みはそれぞれ、購入及びブラウジング振舞いの相対重要度に対応する。

【0041】ステップ520で、複数のクラスタ中心が生成される。図4のステップ420で生成された各仲間グループに対して1つのクラスタ中心が計算される。クラスタ中心は、そのクラスタ(仲間グループ)内の、全ての個々の顧客特徴付けのテキスト特徴付けを連結することにより生成される。ステップ530で、ステップ510で導出された顧客特徴付けに最も近いクラスタ中心が見い出される。顧客特徴付けに“最も近い”クラスタ中心は、当業者には既知の目的関数により達成され得る。“接近度”的決定に有用な目的関数の例には、余弦角、ダイス係数(dice coefficient)、ジャッカード係数(Jaccard coefficient)が含まれる。これに関しては、William B. Frakes及びRicardo Baeza-Yatesによる“Information Retrieval, Data Structure and Algo-

thms”[Prentice Hall, 1992]を参照されたい。そこでは、係数の使用による類似点(すなわち接近度)の定義について述べられている。ステップ540で、そのクラスタ中心が顧客特徴付けに最も近いことが判明したクラスタ(仲間グループ)のメンバに当たる顧客を構成する仲間リストが見い出される。ステップ550で、仲間グループが、特定の電子商取引環境において提起される照会に応答するために使用される。

【0042】使用される照会の例について、図5に関連して述べることにする。次に挙げる照会の各々は、ステップ502で定義された第1の入力を要求するが、特定の照会は、ステップ505で定義された入力のいずれか一方または両方を使用し得る。典型的な照会を以下に示す。

【0043】照会1：購入またはブラウズされた製品のセットに対して、最善の推薦リストを形成する全ての製品のリストを見い出す。

【0044】重み付け連結が、現セッションにおいて顧客により購入またはブラウズされた製品の製品特徴付け(第2の入力)に対して実行される。履歴データからの製品特徴付けを表す第3の入力は、使用されない。最も近い仲間の報告が、ステップ540で生成される。最も近い仲間により購入された、最も高い頻度を有する製品が、次に推薦として報告される。

【0045】照会2：所与の顧客に対して、及びその顧客により購入またはブラウズされた製品のセットに対して、最善の推薦リストを見い出す。

【0046】この照会は、ステップ505の第3の入力の重み付け頻度が、第2の入力の重み付け頻度と共に使用される以外は、照会1に類似である。

【0047】照会3：所与の顧客に対して、最善の推薦リストを見い出す。

【0048】この照会は第1及び第2の照会に類似であるが、ステップ505からの第3の入力だけが使用される。実際、第2の入力が廃棄される。

【0049】照会4：照会1、照会2及び照会3に対して、指定プロモーション・リストから推薦リストを見い出す。

【0050】この照会は、製品推薦リストがフィルタリングされ、ステップ550で指定された指定プロモーション・リスト上に含まれるものを探し、全ての製品を除外するという点で、最初の3つの照会とは異なる。

【0051】照会5：特定の製品に関心を持つ顧客のプロファイルを見い出す。

【0052】この照会は、照会1で述べたように実施される。これは、仲間分布プロファイルがその仲間グループの製品頻度にもとづき出力されるという点で、図5のステップ550とは異なる。仲間分布プロファイルは、年齢、性別などのグループ内の顧客のプロファイルである。例えば、年齢ヒストグラムが出力され得る。

【0053】照会6：顧客の所与のプロファイルに対して、彼らが最も好むであろう製品を見い出す。

【0054】この照会は、次の形式すなわち“20歳乃至30歳のグループの人々は、どのような製品を最も好むであろうか？”といった形式を取る、顧客プロファイル照会を考慮する。この照会は、ステップ510でセットされた指定プロファイル内の全ての顧客のテキスト特徴付けを連結することにより応答される。

【0055】照会7：照会1、照会2及び照会3に対して、製品そのものではないに、顧客に最も好まれるであろう特徴付けを生成する。

【0056】この照会に対して、ステップ520で導出されたクラスタ中心が、要求された製品特徴付けを提供する。

【0057】照会8：所与の顧客に対して、最も近い仲間を見い出す。この照会に対して、ステップ530が要求される応答を提供する。すなわち、選択されたクラスタ中心内に含まれる顧客の報告である。

【0058】有利なことに、本発明による方法及びシステムは、グループ製品の格付けに頼ることなく、顧客の自然なブラウジングまたは購入振舞いにもとづき、顧客及び製品の特徴付けを容易にする。顧客特徴付けは、製品推薦を最終的に生成する手段となり、それにより製品推薦プロセスに本質的な個人化をもたらす。

【0059】本発明の上述の実施例は、付随する図面に関連して述べられてきたが、当業者であれば、本発明がこれらの実施例そのものに限定されるものではなく、本発明の範囲または趣旨から逸れることなく、様々な他の変更及び変形が可能であることが理解できよう。

【0060】まとめとして、本発明の構成に関して以下の事項を開示する。

【0061】(1) 電子商取引環境において、顧客に製品推薦を提供する方法であって、

a) 複数の製品の各々に対する製品特徴付けを導出するステップと、

b) 各それぞれの前記顧客による前記製品特徴付けの利用にもとづき、前記顧客の各々に関する、個々の顧客特徴付けを生成するステップと、

c) 前記顧客特徴付けの類似点にもとづきクラスタ化し、仲間グループを形成するステップと、

d) 個々の前記顧客を前記仲間グループの1つに分類するステップと、

e) 前記顧客特徴付け及び前記分類済み仲間グループからの情報にもとづき、前記顧客に製品推薦を生成するステップとを含む、方法。

(2) 前記顧客特徴付けを生成するステップが、前記顧客が前記製品をブラウズまたは購入する際、前記製品特徴付けを抽出するステップを含む、前記(1)記載の方法。

(3) 前記顧客特徴付けを生成するステップが、個々の

前記顧客によりブラウズまたは購入された全ての前記製品の各前記製品特徴付けを連結するステップを含む、前記(2)記載の方法。

(4) 前記製品特徴付けを導出するステップが、前記複数の製品に関連付けられるテキスト特徴付けを使用する、前記(1)記載の方法。

(5) 前記製品特徴付けを生成するステップが、  
a) 前記テキスト記述内での各単語の発生頻度を見い出すステップと、

b) 各前記単語の合計頻度を、全ての顧客における当該前記単語の発生頻度により除算するステップと、  
c) 各前記単語の標準偏差を見い出すステップと、  
d) より大きな標準偏差を有する単語を選択するステップと、

e) 選択された前記単語にもとづき、製品特徴付けを表現するステップとを含む、前記(1)記載の方法。

(6) a) 現オンライン・セッションにもとづき、前記製品特徴付けを導出するステップと、  
b) メモリから履歴製品特徴付けをアクセスするステップと、

c) ステップa)及びb)からの前記特徴付けを重み付け連結することにより、前記顧客特徴付けを生成するステップと、  
d) 前記仲間グループの各々のクラスタ中心を計算するステップと、

e) ステップc)で生成された前記顧客特徴付けに最も近い前記クラスタ中心を有する前記仲間グループを選択するステップと、  
f) 選択された前記仲間グループにもとづき、製品推薦、仲間グループ推薦及びプロファイル推薦の1つを生成するステップとを含む、前記(1)記載の方法。

(7) 前記クラスタ中心が、各前記仲間グループ内の全ての前記顧客特徴付けのテキスト特徴付けを連結することにより計算される、前記(6)記載の方法。

(8) 前記推薦が、前記現オンライン・セッションにおいて購入及びブラウズされた製品の、テキスト特徴付けの重み付け連結を含む、前記(6)記載の方法。

(9) 電子商取引環境において、顧客に製品推薦を提供するプロセス・ステップを実行する、コンピュータにより実行可能な命令プログラムを実現する、コンピュータにより読み取り可能な記憶プログラム装置であって、前記命令プログラムが、電子商取引において売りに出される製品に関する製品情報をアクセスするステップと、

前記製品情報から製品特徴付けを導出するステップと、前記製品に関する顧客利用情報をアクセスするステップと、

アクセスされた前記利用情報にもとづき、前記顧客の各々に関する顧客特徴付けを生成するステップと、前記顧客特徴付けの類似点にもとづきクラスタ化し、複

数の仲間グループを形成するステップと、現セッション顧客から照会を受信するステップと、前記現セッション顧客に関連付けられる現顧客特徴付けを生成するステップと、前記現セッション顧客を、前記現顧客特徴付け及び過去の顧客特徴付けの類似点にもとづき、選択された前記仲間グループに分類するステップと、前記照会に対して、少なくとも分類された選択済みの前記仲間グループからの情報にもとづき、前記現セッション顧客に製品推薦により応答するステップとを含む、装置。

(10) 前記コンピュータに、前記現セッション顧客と電子商取引サイト間で交換されるデータを抽出し、前記現顧客特徴付けを生成するように指示する記憶プログラムを含む、前記(9)記載の装置。

(11) 前記現顧客特徴付けを生成するステップが、前記現セッション顧客によりブラウズまたは購入された全ての製品の前記製品特徴付けの各々を重み付け連結するステップを含む、前記(9)記載の装置。

(12) 各前記仲間グループ内の全ての前記顧客特徴付けのテキスト特徴付けを連結することにより、当該仲間グループのクラスタ中心を計算するステップと、前記現顧客特徴付けに最も近い前記クラスタ中心を有する前記仲間グループを選択するステップと、選択された前記仲間グループからのデータにもとづき、製品推薦、仲間グループ推薦及びプロファイル推薦の1つを生成するステップとを含む、前記(9)記載の装置。

(13) 電子商取引環境において顧客と対話し、顧客照会に応答を提供するようにコンピュータに指示する、プロセッサ及び記憶プログラムを有するコンピュータであって、該コンピュータが、複数の製品の製品特徴付けを導出する手段と、前記顧客による前記製品特徴付けの利用にもとづき、前記顧客特徴付けを生成する手段と、前記顧客特徴付けの類似点にもとづきクラスタ化し、複数の仲間グループを形成する手段と、個々の前記顧客を前記仲間グループの1つに分類する手段と、前記顧客特徴付け及び前記仲間グループからの情報にもとづき、前記顧客照会に応答を提供する手段とを含む、記憶プログラム。

(14) 前記顧客特徴付けを生成する手段が、前記顧客特徴付けを保管する記憶装置を含む、前記(13)記載の記憶プログラム。

(15) 現オンライン顧客による前記製品特徴付けの利用から抽出されるデータの連結にもとづき、前記現オンライン顧客の現顧客特徴付けを生成する手段を含む、前記(13)記載の記憶プログラム。

(16) 前記現顧客特徴付け及び記憶済みの特徴付けの類似点にもとづき、前記現オンライン顧客を前記仲間グループの1つに分類する手段を含む、前記(15)記載の記憶プログラム。

(17) 前記製品特徴付けを導出する手段が、前記製品に関連付けられるテキスト記述を抽出する手段を含む、前記(13)記載の記憶プログラム。

(18) 前記応答を提供する手段が、現セッションにおいて個々の前記顧客によりブラウズされた製品からの前記製品特徴付けにもとづき、個々の前記顧客に前記製品を推薦する手段を含む、前記(13)記載の記憶プログラム。

(19) 前記応答を提供する手段が、履歴製品特徴付けの重み付け頻度、及び現セッションにおいて個々の前記顧客によりブラウズされた製品からの、前記現製品特徴付けの重み付け頻度にもとづき、個々の前記顧客に前記製品を推薦する手段を含む、前記(13)記載の記憶プログラム。

(20) 前記応答を提供する手段が、履歴製品特徴付けにもとづき、個々の前記顧客に製品を推薦する手段を含む、前記(13)記載の記憶プログラム。

(21) 前記応答を提供する手段が、プロモーション・リスト内の製品にもとづき、個々の前記顧客に製品を推薦する手段を含む、前記(13)記載の記憶プログラム。

(22) 前記応答を提供する手段が、前記仲間グループの1つ内の製品頻度にもとづき、仲間分布プロファイルを提供する手段を含む、前記(13)記載の記憶プログラム。

### 30 【図面の簡単な説明】

【図1】本発明に従うシステムのブロック図である。

【図2】本発明に従うシステムの全体動作の例を示すフロー図である。

【図3】本発明に従い、製品特徴付けを生成する方法を示すフロー図である。

【図4】本発明に従い、顧客特徴付けを生成する方法を示すフロー図である。

【図5】本発明に従い、推薦を生成する方法を示すフロー図である。

### 40 【符号の説明】

4 顧客またはクライアント

6 ネットワーク

7 要求(照会)

8 サーバ

9 照会結果

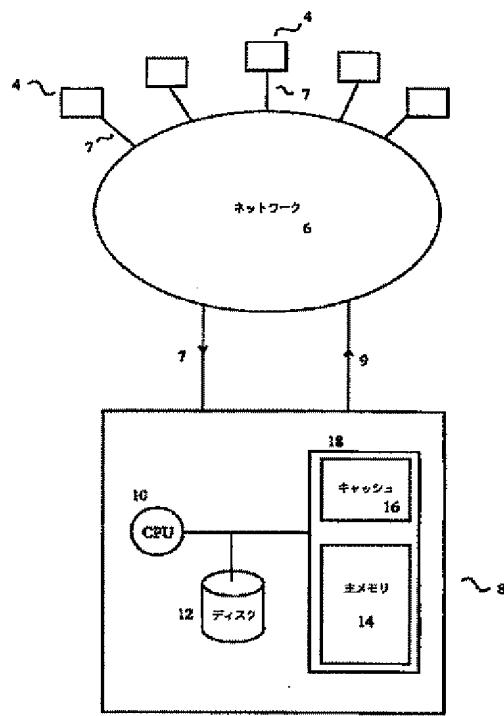
10 CPU

12 ディスク

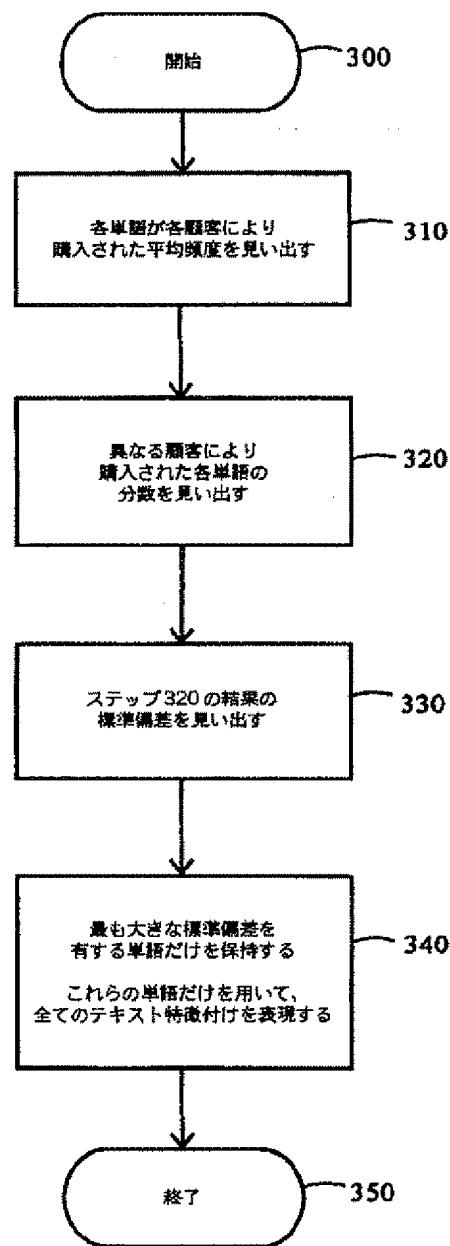
14 主メモリ

16 キャッシュ

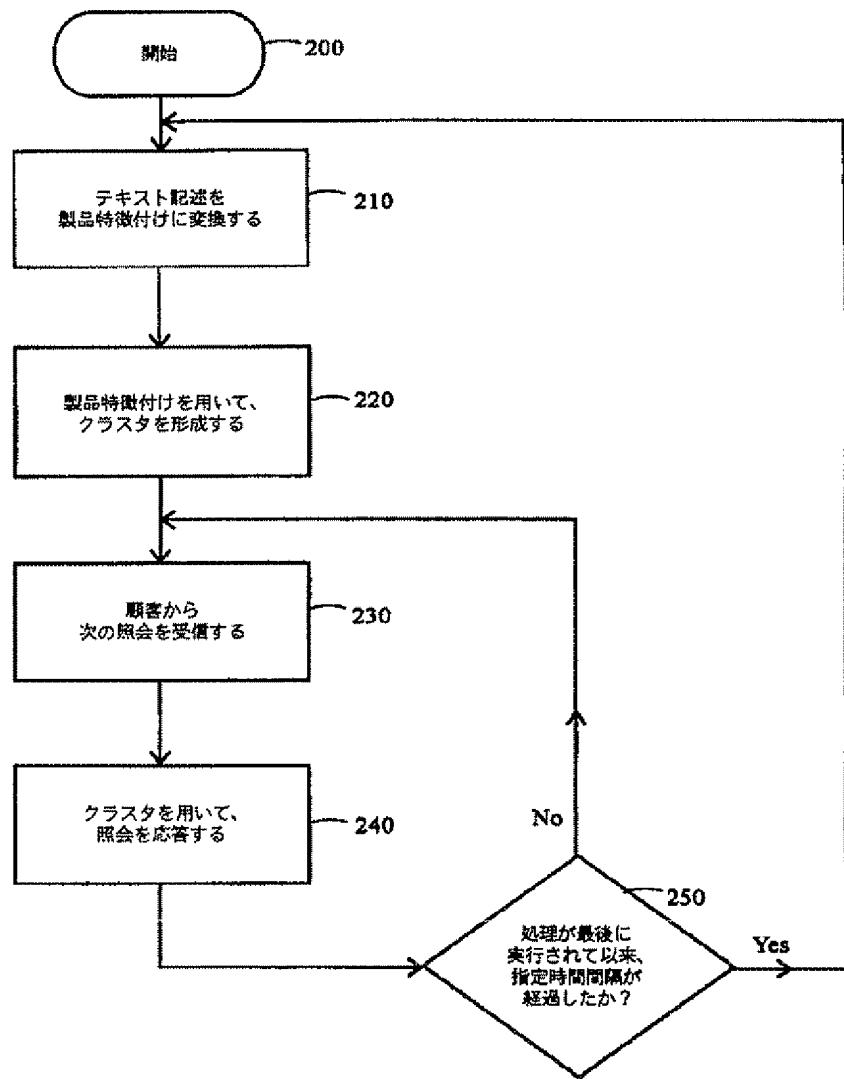
【図1】



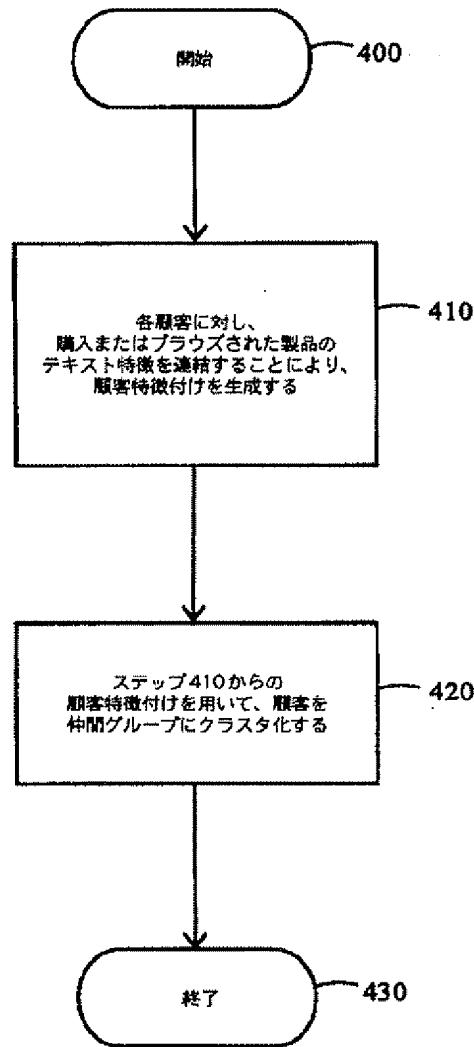
【図3】



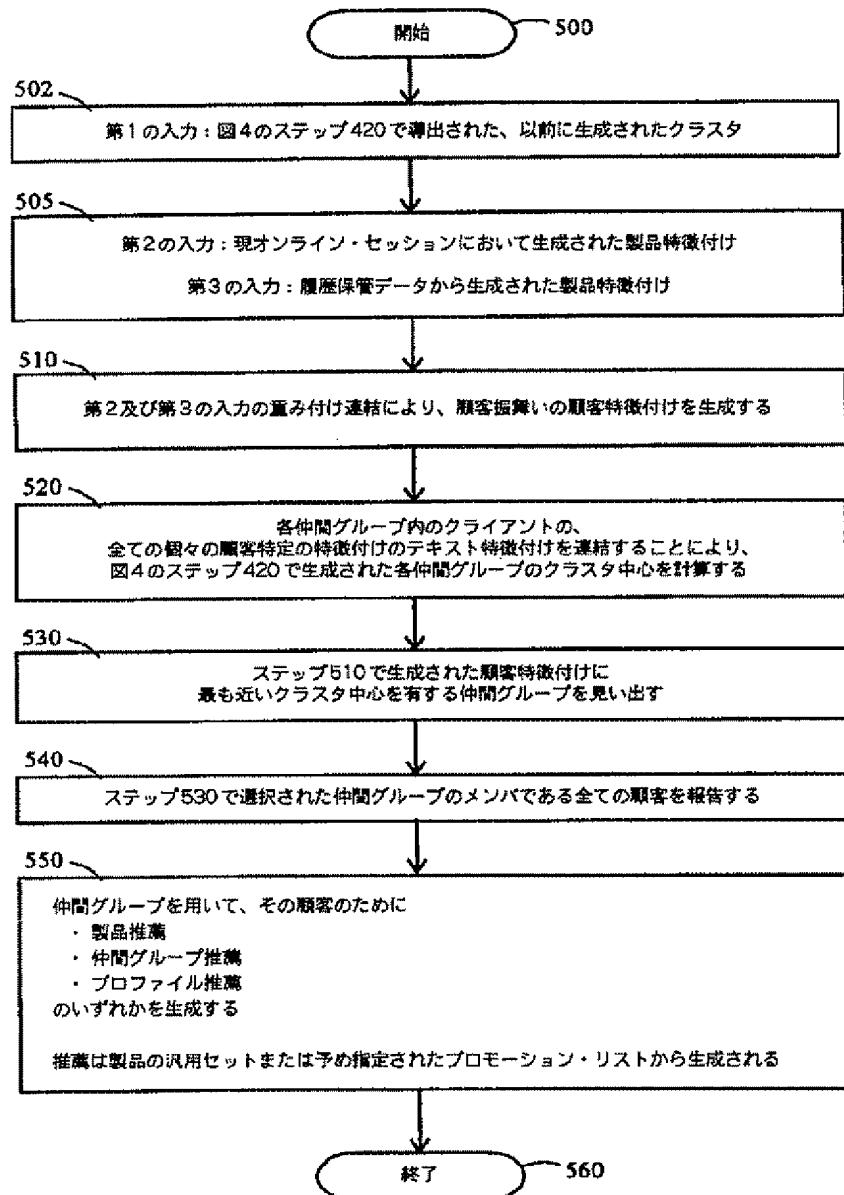
【図2】



【図4】



【図5】



フロントページの続き

(72)発明者 チャル・シィ・アガーワル  
アメリカ合衆国10598、ニューヨーク州ヨ  
ークタウン・ハイツ、コンパウンド・ロー  
ド 2219

(72)発明者 フィリップ・エス・ユ  
アメリカ合衆国10514、ニューヨーク州チ  
ャバクラ、ストーンウェイ 18

